Reconstruction of positive solutions for ill-posed problems

Elena Resmerita

Institute of Mathematics
Alpen-Adria Universität
Klagenfurt, Austria

Splitting Algorithms, Modern Operator Theory, and Applications
September 17-22, 2017
CMO-BIRS, Oaxaca, Mexico

Introduction

Consider an ill-posed linear operator equation

$$Au = y$$

with $A : L^1(\Omega) \to Y$ bounded, where Ω a bounded and closed subset of $\mathbb{R}^d$, and Y is a separable Hilbert space.

Aim: Recovering stably a nonnegative solution of the equation, when it exists.

Presentation based on

- a survey co-authored with
  - Barbara Kaltenbacher, Klagenfurt University
  - Christian Clason, Duisburg-Essen University
- a joint work with Martin Burger, Münster University.

Entropy functionals

The (negative of the) Boltzmann-Shannon entropy $f : L^1(\Omega) \to (-\infty, +\infty]$ is defined as

$$f(u) = \begin{cases} \int_\Omega u(t) \log u(t)\, dt & \text{if } u \geq 0 \text{ a.e. and } u \log u \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

The Kullback-Leibler functional $d : dom\, f \times dom\, f \to [0, +\infty]$ is

$$d(v,u) = f(v) - f(u) - f'(u, v-u),$$

$$d(v,u) = \int_\Omega \left[v(t) \ln \frac{v(t)}{u(t)} - v(t) + u(t) \right] dt, \qquad (1)$$

when it is finite.

Useful properties of the entropy functionals

- The function f is strictly convex and lower semicontinuous with respect to the (weak) topology of $L^1(\Omega)$.
- For any $c > 0$, the sublevel set

$$\left\{v \in L^1_+(\Omega) : f(v) \leq c\right\}$$

  is convex, (weakly) closed, and weakly compact in $L^1(\Omega)$.
- The interior of the domain of the function f is empty.
- The set $\partial f(u)$ is nonempty if and only if u belongs to $L^\infty_+(\Omega)$ and is bounded away from zero. In this case, $\partial f(u) = \{1 + \log u\}$.

Borwein and Lewis '91, Amato and Hugh '91, Borwein and Limber '96

Variational methods for recovering nonnegative solutions

Let $Au = y$ and let y^δ be the noisy data satisfying $\|y^\delta - y\| \leq \delta$ with $\delta > 0$.

Denote by u_0 some a priori guess of the solution.

- Maximum entropy regularization

$$\min_{u \geq 0} \|Au - y^\delta\|^2 + \alpha \mathcal{R}(u),$$

  for some regularization parameter $\alpha > 0$, where $\mathcal{R} \in \{f, d(\cdot, u_0)\}$.

- Denote u_α^δ the (unique) solution of the above problem.

Computationally: nonlinear optimization problems.

Engl, Landl, Eggermont

Convergence results

Let $\alpha = \alpha(\delta)$ be chosen such that

$$\alpha \to 0 \text{ and } \frac{\delta^2}{\alpha} \to 0 \text{ as } \delta \to 0\,.$$

Then the minimizers u_α^δ converge to the maximum entropy solution $u^\dagger$ of $Au = y$ (that is, $u^\dagger = \arg\min d(u, u_0)$ s.t. $Au = y$):

$$\|u_\alpha^\delta - u^\dagger\|_1 \to 0 \text{ as } \delta \to 0\,.$$

Moreover, if $u^\dagger$ satisfies the (source) condition

$$\log \frac{u^\dagger}{u_0} = A^* w$$

for some $w \in Y$, and $\alpha = \alpha(\delta)$ is chosen such that $\alpha \sim \delta$ as $\delta \to 0$, then one has

$$\|u_\alpha^\delta - u^\dagger\|_1 = \mathcal{O}(\sqrt{\delta}).$$

More details

- The following inequalities are essential to obtaining the previous results:

$$\begin{aligned} \tfrac{1}{2}\|A(u-u_\alpha^\delta)\|_Y^2 + \alpha d(u,u_\alpha^\delta) &\leq \tfrac{1}{2}\|Au-y^\delta\|_Y^2 + \alpha d(u,u_0) \\ &- \tfrac{1}{2}\|Au_\alpha^\delta - y^\delta\|_Y^2 - \alpha d(u_\alpha^\delta, u_0), \quad \forall u \end{aligned}$$

  and

$$\tfrac{1}{2}\|A(\tilde{u}_\alpha^\delta - u_\alpha^\delta)\|_Y^2 + \alpha d(\tilde{u}_\alpha^\delta, u_\alpha^\delta) \leq 2\|\tilde{y}^\delta - y^\delta\|_Y^2,$$

  where $\tilde{u}_\alpha^\delta$ is the minimizer corresponding to $\tilde{y}^\delta$ instead of y^δ.

- Uniform positivity of u_α^δ is obtained, that is $\frac{u_\alpha^\delta}{u_0}$ is bounded away from zero.

  Eggermont '93

Versions of entropy regularization

- Morozov-entropy regularization:

$$\min_{u\in L^1_+(\Omega)} d(u,u_0) \quad \text{s.t.} \quad \|Au-y^\delta\|\le\delta.$$

No regularization parameter α!

Amato and Hugh '91

- Ivanov-entropy regularization (method of quasi-solutions)

$$\min_{u\in L^1_+(\Omega)} \|Au-y^\delta\| \quad \text{s.t.} \quad f(u)\le\rho$$

ρ is the regularization parameter. Ivanov '62

- Tikhonov-entropy regularization (presented before);

Another approach:

$$\min_{u\in\mathcal{D}} \frac{1}{2}\|Au-y^\delta\|^2+\alpha f(u)$$

The analysis relies on a nonlinear transformation T with $f(T(v))=\|v\|_2^2+c$, where $T:\left\{v\in L^2(\Omega):v\ge c, \text{ a.e.}\right\}\to L^1(\Omega)$. Engl and Landl '92

Versions of entropy regularization

- If the respective minimizers are unique, all three variational regularization methods are equivalent for a certain choice of the regularization parameters α and ρ.
- A practically relevant regularization parameter choice might lead to different solutions.
- The three formulations also entail different numerical approaches, some of which might be better suited than others in concrete applications.

Interesting: Better understanding of the solutions of the three entropy methods.

Lorenz and Worliczek '13

Joint Kullback-Leibler regularization:

$$\min_{u \geq 0} d(y^\delta, Au) + \alpha d(u, u_0), \quad \alpha > 0,$$

for problems where A has positive values and y is positive.

R. and Anderssen '08

Iterative regularization methods for positive solution reconstruction

- Iterative methods for ill-posed problems have a typical behavior:
  The distance between the solution $u^\dagger$ and the iterates u_k^δ decays initially, then it increases.

- It is necessary to choose an appropriate stopping index $k_* := k_*(\delta, y^\delta) < \infty$ such that $u_{k_*}^\delta \to u^\dagger$ as $\delta \to 0$.
- A frequent choice is a discrepancy principle, e.g., of Morozov.

Iterative regularization methods for positive solution reconstruction

- Iterative methods for ill-posed problems have a typical behavior:
  The distance between the solution $u^\dagger$ and the iterates u_k^δ decays initially, then it increases.

- It is necessary to choose an appropriate stopping index $k_* := k_*(\delta, y^\delta) < \infty$ such that $u_{k_*}^\delta \to u^\dagger$ as $\delta \to 0$.
- A frequent choice is a discrepancy principle, e.g., of Morozov.

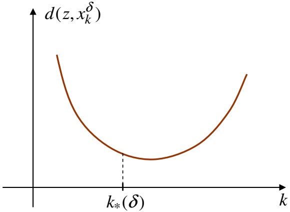
Iterative regularization methods for positive solution reconstruction

- Iterative methods for ill-posed problems have a typical behavior:
  The distance between the solution $u^\dagger$ and the iterates u_k^δ decays initially, then it increases. Engl, Hanke, Neubauer '96

- It is necessary to choose an appropriate stopping index $k_* := k_*(\delta, y^\delta) < \infty$ such that $u_{k_*}^\delta \to u^\dagger$ as $\delta \to 0$.
- A frequent choice is a discrepancy principle, e.g., of Morozov.

Iterative regularization methods for positive solution reconstruction

- Iterative methods for ill-posed problems have a typical behavior:
  The distance between the solution $u^\dagger$ and the iterates u_k^δ decays initially, then it increases. Engl, Hanke, Neubauer '96



- It is necessary to choose an appropriate stopping index $k_* := k_*(\delta, y^\delta) < \infty$ such that $u_{k_*}^\delta \to u^\dagger$ as $\delta \to 0$.
- A frequent choice is a discrepancy principle, e.g., of Morozov.

Landweber method in Hilbert spaces

Classical Landweber method:
Choose $u_0 = 0$, $\tau \in (0, 2\|A\|^{-2})$ and

$$u_{k+1} = u_k + \tau A^*(y - Au_k), \qquad k = 0, \ldots.$$

- The iterates converge strongly to the minimum norm solution $u^\dagger$ for exact data $y \in \operatorname{ran} A$.
- For noisy data $y = y^\delta \in \overline{\operatorname{ran} A} \setminus \operatorname{ran} A$: convergence when using the discrepancy principle of Morozov:
  $k_*(\delta, y^\delta) = \max\{k \in \mathbb{N} : \|Au_k - y^\delta\| \geq \tau\delta\}$, for some $\tau > 1$.

  Remark: Stability estimate of order $\mathcal{O}(\sqrt{k})$,

$$\|u_k^\delta - u_k\| = \mathcal{O}(\sqrt{k}).$$

Projected Landweber methods in Hilbert spaces

- Generalization to constrained inverse problems of the form

$$Au = y \quad \text{s.t.} \quad u \in C$$

for a convex and closed set $C \subset L^2(\Omega)$.
The projected Landweber method:

$$u_{k+1} = P_C \left[u_k + \tau A^* (y - Au_k) \right], \qquad k = 0, \ldots,$$

where P_C is the metric projection onto C.
Eicke '92

- This coincides with a forward–backward splitting or proximal gradient descent applied to $\|Au - y\|^2 + \delta_C(u)$, where δ_C is the indicator function of C.
Combettes, Wajs '05

Projected Landweber methods in Hilbert spaces

Convergence results for projected Landweber:

- The case of exact data $y \in A(C)$:
  - Weak convergence of the iterates to $u^\dagger$.
  - Strong convergence under an additional restrictive condition: $\mathrm{Id} - \tau A^* A$ is compact.
- The case of noisy data $y^\delta \notin \operatorname{ran} A$:.

$$\|u_k^\delta - u_k\| \le \tau \|A\| \delta k, \qquad k = 0, \ldots.$$

"Dual" projected Landweber iteration:

- Set $w_0 = 0$, compute for $k = 0, \ldots,$ the iterates

$$\begin{cases} u_k = P_C A^* w_k, \\ w_{k+1} = w_k + \tau(y - Au_k). \end{cases}$$

- Strong convergence can be shown; preconditioning also possible (Viana '97)
  Hope: Further acceleration by adding inertial terms.

Expectation-Maximization algorithms for integral equations

- Consider Fredholm integral operators of the first kind, i.e.,

$$A : L^1(\Omega) \to L^1(\Sigma), \qquad (Au)(s) = \int_\Omega a(s,t)u(t)\,dt\,, \quad (2)$$

  where the kernel a and the data y are positive pointwise a.e.
- The method of convergent weights:

$$u_{k+1}(t) = u_k(t) \int_\Sigma \frac{a(s,t)y(s)}{(Au_k)(s)}\,ds, \quad t \in \Omega, \qquad k = 0, \dots .$$

Kondor '83

The finite dimensional counterpart

$$u_{k+1}^j = u_k^j \sum_{i=1}^{n} \frac{a_{ij} y_i}{\sum_{l=1}^{m} a_{il} u_k^l}, \quad j = 1, m,$$

- EM algorithm for PET and the Lucy–Richardson algorithm in astronomical imaging.
  Richardson '72, Lucy '75, Shepp and Vardi '82, Iusem '92

  Bausche, Noll, Celler, Borwein '91
- It converges to (nonnegative) minimizers of $d(y, Au)$.
- Advantages of EM:
  - it shapes the features of the solution in early iterations
  - easy to compute
- Disadvantages of EM:
  - slow algorithm
  - very unstable numerically

Properties of the EM algorithm in infinite dimension

- Under some assumptions on A, y, one can show:

  If $u_0 \in L^1_+(\Omega)$ such that $d(u^\dagger, u_0) < \infty$, then, for any $k \geq 0$, the iterates u_k satisfy

$$d(u^\dagger, u_k) < \infty,$$

$$d(u_{k+1}, u_k) \leq d(y, Au_k) - d(y, Au_{k+1}),$$

$$d(y, Au_k) - d(y, Au^\dagger) \leq d(u^\dagger, u_k) - d(u^\dagger, u_{k+1}).$$

  Therefore, the sequences $\{d(u^\dagger, u_k)\}_{k\in\mathbb{N}}$ and $\{d(y, Au_k)\}_{k\in\mathbb{N}}$ are nonincreasing. Moreover,

$$\lim_{k\to\infty} d(y, Au_k) = d(y, Au^\dagger),$$

$$\lim_{k\to\infty} d(u_{k+1}, u_k) = 0.$$

Mülthei and Schorr, Eggermont

Open problem: Convergence of the algorithm.

The EM algorithm with noisy data y^δ satisfying $d(y^\delta, y) \le \delta$

$$u_{k+1}^\delta(t) = u_k^\delta(t) \int_\Sigma \frac{a(s,t) y^\delta(s)}{(Au_k)(s)} ds, \quad t \in \Omega, \qquad k = 0, \ldots .$$

- One can show:

$$d(u^\dagger, u_{k+1}^\delta) \le d(u^\dagger, u_k^\delta)$$

for all $k \ge 0$ such that

$$d(y^\delta, Au_k^\delta) \ge \delta\gamma,$$

for some constant $\gamma > 0$.

- A possible choice of the stopping index for the algorithm:

$$k_*(\delta) = \min\left\{k \in \mathbb{N} : d(y^\delta, Au_k^\delta) \le \tau\delta\gamma\right\}$$

for some fixed $\tau > 1$.

The EM algorithm with noisy data y^δ satisfying $d(y^\delta, y) \le \delta$

- A stopping index for the algorithm:

$$k_*(\delta) = \min\left\{k \in \mathbb{N} : d(y^\delta, Au_k^\delta) \le \tau\delta\gamma\right\}$$

for some fixed $\tau > 1$.

- Existence of such a stopping index:
For all $\delta > 0$, there exists a $k_*(\delta)$ defined above such that

$$k_*(\delta)\tau\delta\gamma \le k_*(\delta) d(y^\delta, Au^\delta_{k_*(\delta)-1}) \le d(u^\dagger, u_0) + k_*(\delta)\delta\gamma.$$

The stopping index $k_*(\delta)$ is finite:

$$k_*(\delta) \le \frac{d(u^\dagger, u_0)}{\gamma(\tau - 1)\delta}$$

and

$$\lim_{\delta \to 0^+} \|Au^\delta_{k_*(\delta)} - y\|_p = 0,$$

for any $p \in [1, +\infty)$. R, Engl, Iusem '07

More about EM in infinite dimension

- EM algorithms with smoothing steps:

$$u_{k+1} = S\left(\mathcal{N}(u_k)\, A^* \frac{y}{Au_k}\right), \qquad k = 0, \dots, \tag{3}$$

with $u_0 \equiv 1$ and

$$\mathcal{N}u(t) = \exp\left([S^*(\log u)](t)\right) \qquad \text{for all } t \in \Omega. \tag{4}$$

Here S is a linear smoothing (integral) operator.

Eggermont '96

- OS-EM method:

Haltmeier, Leitao, R '09

Interesting: Convergence of the algorithm by using the discrepancy principle or other stopping rules.

Entropic projection method in infinite dimensional spaces

$$u_k \in \arg\min_u \left\{ \frac{1}{2}\|Au - y^\delta\|^2 + \mu d(u, u_{k-1}) + \chi_j(u) - \frac{1}{2}\|Au - Au_{k-1}\|^2 \right\},$$

equivalently,

$$u_k \in \arg\min_u \left\{ \langle Au, Au_{k-1} - y^\delta \rangle + \mu d(u, u_{k-1}) + \chi_j(u) \right\},$$

where

$$\chi_1(u) = \begin{cases} 0 & \text{if } \int_\Omega u(t)\ dt = 1, \\ +\infty & \text{else,} \end{cases}$$

and $\chi_0 \equiv 0$ (the original problem without integral constraint), $\mu > 0$.

- One can show welldefinedness of the iterates u_k.
- Nonnegativity: u_0 nonnegative $\Rightarrow$ u_k nonnegative, $\forall k \in N$.

Yin, Osher, Goldfarb and Darbon '08, Burger, R '17

The theoretical context

We work with operators satisfying a 'continuity' condition:

$$\|Au - Av\| \leq \gamma\sqrt{d(u,v)} \text{ for some } \gamma > 0 \quad (cc)$$

The two situations we consider:

- Mean one constraint, that is $\int u_k(t)\,dt = 1, k \in \mathbb{N}$:

$$u_k = c_{k-1} u_{k-1} e^{\lambda A^*(y^\delta - Au_{k-1})}, \qquad c_{k-1} = \frac{1}{\int_\Omega u_{k-1} e^{\lambda A^*(y^\delta - Au_{k-1})}\, dt},$$

  $\surd$ (cc)

- No mean constraint;

$$u_k = u_{k-1} e^{\lambda A^*(y^\delta - Au_{k-1})},$$

with $\lambda = 1/\mu$ (pointwise equalities defining u_k).

Examples of operators satisfying (cc)?

Related literature in finite dimensional optimization

$$u_k \in \arg\min_u \{\langle u, A^*(Au_{k-1} - y)\rangle + \mu_k d(u, u_{k-1})\},$$

with $d = D_f = KL$, f being the entropy: $f(u) = \sum_{j=1}^n u_j \ln u_j$.

Start with:

$$\min_{u \geq 0} g(u)$$

- Proximal point methods:

$$u_{k+1} = \text{argmin}_u g(u) + \mu_k d(u, u_k)$$

  Implicite iterative method

- Easier: Linearize the objective functional, i.e., $g(u) \sim g(u_k) + \nabla g(u_k)^t (u - u_k)$

$$u_{k+1} = \text{argmin}_u \nabla g(u_k)^t u + \mu_k d(u, u_k)$$

$$u_{k+1} = \text{argmin}_u \nabla g(u_k)^t u + \mu_k d(u, u_k)$$

- The first order optimality condition for this problem is

$$\nabla f(u_{k+1}) = \nabla f(u_k) - \frac{1}{\mu_k} \nabla g(u_k),$$

Since ∇f invertible,

$$u_{k+1} = (\nabla f)^{-1}(\nabla f(u_k) - \frac{1}{\mu_k} \nabla g(u_k))$$

that is

$$u_{k+1}^j = u_k^j e^{-\lambda_k \nabla g(u_k)^j}, \quad \lambda_k = 1/\mu_k$$

- Several line search versions of the algorithm have been proposed and analyzed.

Iusem '94, '97

Convergence analysis - Noisy data case
Discrepancy principle

Proposition: If

- $A : L^1(\Omega) \to Y$ is bounded and linear, satisfying the 'continuity' condition
- z is a positive solution of $Au = y$ with $\chi_j(z) = 0$ if $j = 1$.
- $y^\delta \in Y$ are noisy data satisfying $\|y - y^\delta\| \leq \delta$, for some noise level δ
- $u_0 \in dom\,\partial f$ is an arbitrary starting element with the properties $1 + \log u_0 \in \mathcal{R}(A^*)$ and $\chi_j(u_0) = 0$ if $j = 1$
- the stopping index k_* is chosen such that

$$k_*(\delta) = \min\{k \in \mathbb{N} : \|Au_k - y^\delta\| \leq \tau\delta\}, \ \ \tau > 1,$$

then

i) The residual $\|Au_k - y^\delta\|$ decreases when k increases.
ii) The index $k_*(\delta)$ is finite.
iii) There exists a weakly convergent subsequence of $(u_{k_*(\delta)})_\delta$ in $L^1(\Omega)$. If $(k_*(\delta))_\delta$ is unbounded, then each limit point is a solution of $Au = y$.

Convergence analysis - Noisy data case II
A priori rule

Proposition: If

- $A : L^1(\Omega) \to Y$ is bounded and linear, satisfying the 'continuity' condition
- z is a positive solution of $Au = y$ with $\chi_j(z) = 0$ if $j = 1$.
- $y^\delta \in Y$ are noisy data satisfying $\|y - y^\delta\| \le \delta$, for some noise level δ
- $u_0 \in dom\,\partial f$ is an arbitrary starting element with the properties $1 + \log u_0 \in \mathcal{R}(A^*)$ and $\chi_j(u_0) = 0$ if $j = 1$.
- the stopping index k_* us chosen of order $1/\delta$,

then the sequence $(f(u_{k_*(\delta)}))_\delta$ is bounded and thus, there exists a subsequence of $(u_{k_*(\delta)})_\delta$ in $L^1(\Omega)$ which converges weakly to a solution of $Au = y$.

Error estimates - exact data case

Proposition: If

- $A : L^1(\Omega) \to Y$ is a bounded linear operator satisfying the 'continuity' condition
- z is a positive solution of $Au = y$ verifying $\chi_j(z) = 0$ if $j = 1$
- $u_0 \in dom\, \partial f$ be an arbitrary starting element with the properties $1 + \log u_0 \in \mathcal{R}(A^*)$ and $\chi_j(u_0) = 0$ if $j = 1$.

Additionally, let the following source condition hold:

$$1 + \log z \in \mathcal{R}(A^*).$$

Then one has

$$d(z, u_k) = O(1/k).$$

Moreover, $\|u_k - z\|_1 = O(1/\sqrt{k})$ if $j = 1$.

References

- U. Amato and W. Hughes, Maximum entropy regularization of Fredholm integral equations of the first kind, Inv. Probl. 7 (1991) 793–803.
- M. Burger and E. Resmerita, Iterative Regularization of Linear Ill-posed Problem by an Entropic Projection Method, preprint, 2017.
- P. Eggermont. Maximum entropy regularization for Fredholm integral equations of the first kind. SIAM J. Math. Anal. 24 (1993) 1557–1576.
- H. Engl and G. Landl. Convergence-rates for maximum-entropy regularization. SIAM J. Numer. Anal. 30 (1993) 1509-1536.
- A. Kondor. Method of convergent weights An iterative procedure for solving Fredholm's integral equations of the first kind. Nucl. Instr. Meth. Phys. Res. 216 (1983) 177–181.
- H.N. Mülthei and B. Schorr. On an iterative method for a class of integral equations of the first kind. Math. Methods Appl. Sci. 9(2) (1987) 137–168.
- E. Resmerita and R.S. Anderssen. A joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems. Math. Meth. Appl. Sci. DOI: 10.1002/mma.855 2007.